Steven P. Lehotsky*
Scott A. Keller*
Jeremy Evan Maltz*
LEHOTSKY KELLER COHN LLP
200 Massachusetts Avenue, NW, Suite 700
Washington, DC 20001
(512) 693-8350
steve@lkcfirm.com
scott@lkcfirm.com
jeremy@lkcfirm.com

Joshua P. Morrow*
LEHOTSKY KELLER COHN LLP
408 W. 11th Street, 5th Floor
Austin, TX 78701
(512) 693-8350
josh@lkcfirm.com

* *Pro hac vice* forthcoming.

Bradley A. Benbrook (SBN 177786)
Stephen M. Duvernay (SBN 250957)
BENBROOK LAW GROUP, PC
701 University Avenue, Suite 106
Sacramento, CA 95825
Telephone: (916) 447-4900
brad@benbrooklawgroup.com
steve@benbrooklawgroup.com

Attorneys for Plaintiff NetChoice

UNITED STATES DISTRICT COURT
EASTERN DISTRICT OF CALIFORNIA
SACRAMENTO DIVISION

CHRISTOPHER KOHLS, *et al.*,

Plaintiffs,

v.

ROBERT A. BONTA, *et al.*,

Defendants.

Case No. 2:24-cv-02527-JAM-CKD

NETCHOICE'S PROPOSED AMICUS CURIAE BRIEF IN SUPPORT OF PLAINTIFFS' MOTION FOR SUMMARY JUDGMENT

Judge: Hon. John A. Mendez

TABLE OF CONTENTS

TABLE OF AUTHORITIES

Page(s)

Cases

Statutes

Other Authorities

IDENTITY AND INTEREST OF AMICUS CURIAE

Amicus curiae NetChoice is a national trade association of e-commerce and online businesses that share the goal of promoting convenience, choice, and commerce on the internet. For over a decade, NetChoice has worked to increase consumer access and options via the internet, while minimizing burdens on small businesses that are making the internet more accessible and useful. A full list of NetChoice's members is included here: https://perma.cc/6AXT-6PWG.[1]

Among other things, NetChoice advocates on behalf of its membership for free expression and free enterprise on the Internet, participating in litigation involving issues of vital concern to the online community. Particularly relevant here, NetChoice litigates to vindicate its members' right to moderate the content on their services free from governmental interference. *See, e.g.*, *Moody v. NetChoice, LLC*, 603 U.S. 707 (2024); *NetChoice, LLC v. Fitch*, 738 F. Supp. 3d 753 (S.D. Miss. 2024).

This Court's determinations regarding California Assembly Bill No. 2655's ("AB2655") removal, labeling, and reporting requirements for political speech online will have a significant impact on online businesses whose services are accessed by individuals located in California. The law poses a grave First Amendment risk to both those services and their users. NetChoice is well-situated to explain both AB2655's legal flaws and the practical difficulties that compliance with the law will impose on online services.

[1]Plaintiff X Corp. is a member of NetChoice. But no counsel for a party authored this brief in whole or in part, and no party made a monetary contribution intended to fund the preparation or submission of this brief.

INTRODUCTION & SUMMARY OF THE ARGUMENT

AB2655 requires some of the Internet's largest websites to remove or label certain content-based categories of protected political speech, which the State has dubbed "materially deceptive." §§ 20513-15.[2] But States cannot regulate "the content choices the major platforms make . . . free of the First Amendment's restraints." *Moody*, 603 U.S. at 726-27. Nor may they "deputize[] covered businesses into serving as censors for the State." *NetChoice, LLC v. Bonta*, 113 F.4th 1101, 1118 (9th Cir. 2024). That is why this Court enjoined enforcement of a similar California law. *Kohls v. Bonta*, 2024 WL 4374134, at *4 (E.D. Cal. Oct. 2, 2024). The Court should do the same here by granting Plaintiffs' motion for summary judgment and enjoining Defendants' enforcement of AB2655 against all regulated services.

As this court has already found, the Act would unconstitutionally replace websites' *private* content-moderation decisions with government compulsion to censor political speech. AB2655 singles out websites that disseminate a "staggering amount of content" created by their users. *Moody*, 603 U.S. at 719. In so doing, it targets websites that allow for speech and counter-speech on myriad topics. Yet, those websites do not disseminate *all* user-generated content. They have vigorous "content moderation" policies—that is, efforts "to filter, prioritize, and label the varied messages, videos, and other content their users wish to post." *Id.* at 717. These content-moderation "choices rest on a set of beliefs about which messages are appropriate," "they give the feed[s]" on the services and other aspects of these services "a particular expressive quality," and they "'constitute the exercise'" of protected "'editorial control.'" *Id.* at 738. Among other things, many websites' content-moderation policies govern "misinformation," political satire, parody, and other forms of criticism. *See, e.g.*, *id.* at 719-20.

Content moderation can be difficult. So governmental requirements to censor disfavored political speech risk trampling over the expressive rights of both websites and their users. The massive scale of regulated websites means that those websites must review millions, or even billions, of posts every day. To determine whether any individual piece of content violates websites'

[2] This brief collectively refers to any "public-facing internet website, web application, or digital application," § 20512(h), as "websites" or "services." Statutory citations in this brief refer to the California Elections Code.

private policies requires nuanced, fact-specific judgments about context, intent, and meaning. Making these determinations about *political* speech is especially difficult. Political speech is often deliberately provocative, exaggerated, or ambiguous. The potential for manipulation or error in moderating political speech at the government's behest is greatly increased, as moderators risk inadvertently suppressing speech that is critical or satirical, yet constitutionally protected from government incursion.

Consequently, AB2655 would make content moderation exponentially more difficult by replacing individual websites' unique and ever-evolving policies with the State's one-size-fits-all solution of removal and labeling. The Act also would impose vague governmental demands to determine what political content is, for instance, "reasonably likely" to "undermine confidence" in an election's outcome or "harm the reputation or electoral prospects of a candidate"—as long as it is not "parody" and "satire." §§ 20513(a)(2), 20519. Websites that guess incorrectly face countless lawsuits asking the courts to ensure that websites' editorial practices match the State's mandates. In short, the Act "grants the State nearly unfettered discretion" to supervise websites' content moderation decisions. *NetChoice, LLC v. Bonta*, 2025 WL 807961, at *23 (N.D. Cal. Mar. 13, 2025) ("*Bonta II*"). This will chill the dissemination of speech by the websites. *Id.* at *22.

Federal law prohibits this. Under the First Amendment, States cannot force websites to adopt government-mandated content-moderation policies or enforce state-preferred speech rules. *See, e.g.*, *Moody*, 603 U.S. at 726-27. And under 47 U.S.C. § 230(c), States cannot impose liability on websites for the decisions they make about whether and how to disseminate third-party speech.

Plaintiffs' motion for summary judgment should be granted.

ARGUMENT

I. AB2655's requirements for regulated websites to label and remove certain forms of political speech are unnecessary and unworkable.

A. Regulated websites provide their users with the opportunity to engage in a vast array of political speech and counter-speech.

People use social media and similar websites to access and discuss a wide array of political content. This speech entails text, images, audio, and video from other voters, candidates, elections officials, political organizations, and news outlets. The prevalence of such content is part of what

makes these websites useful outlets for political expression, voter education, and democratic participation. Yet AB2655 regulates these websites by requiring them to identify, remove, and label so-called "materially deceptive content" involving "candidate[s] for elective office," "elections official[s]," and "elected officials." § 20513(a)(2). By doing that, AB2655 inevitably interferes with users' ability to engage in and access protected and valuable political speech on those topics.

For example, voters use these websites to directly engage in the political process. They connect with like-minded citizens, organize grassroots campaigns, and coordinate political activities like voter registration drives or get-out-the-vote efforts. These websites also help all users—and especially young and previously disengaged voters—to more easily participate in political processes. This lowers the barrier to civic engagement. Meanwhile, these websites expose users to opinions and discussions from beyond their immediate social spheres.

Similarly, candidates "for elective office" use social media. *See* § 20513(a)(2)(A). These websites allow them to communicate directly with voters, articulate policy positions, and quickly respond to an opponent's claims. Candidates use these websites to share campaign content, recruit volunteers, organize events, and solicit contributions as well. This offers cost-effective methods for candidates to spread their message. In the same vein, candidates can humanize themselves by sharing authentic and relatable content that allows them to connect personally with voters. Voters, in turn, may benefit from this direct candidate interaction. By following candidates online, voters receive real-time updates about campaigns, electoral developments, and policies.

Next, by allowing users to discuss "elected official[s]," § 20513(a)(2)(C), office holders, and politicians in general, social media websites enable constituents to share information about officials' voting records, policy positions, public statements, and performance in office. This creates ongoing dialogue about how representatives are serving their communities. Social media allows users to discuss legislative decisions, policy initiatives, and official actions taken by elected representatives. All of this helps voters stay informed about their representatives' activities and engage with local issues. By using these websites, communities can have substantive discussions about how elected officials are fulfilling their campaign promises and serving their constituents.

Likewise, "elections official[s]" use social media to facilitate dialogue about election administration. § 20513(a)(2)(B). For example, voters can share their experiences at polling locations, discuss procedures, and provide feedback to election authorities. These conversations educate communities on election processes, enhance transparency, and allow officials to improve voter experiences. Similarly, "elected official[s]" supervising elections benefit from robust online discussion of their oversight roles. § 20513(a)(2)(C).

Some of this valuable political content can involve digitally created or modified audio or visual media. Users rely on digital content to better understand candidates and elections, such as edited videos highlighting candidate messages, interactive visualizations of campaign finance data, digital timelines illustrating voting records, or video summaries of public events. Such media helps voters evaluate candidates and policies. Similarly, some digitally modified content explicitly portrays candidates, election officials, or elected representatives saying or doing things they did not literally say or do, but in helpful ways. For instance, computer-generated videos might show candidates speaking clearly in voters' native languages or might convert written statements into audio form for greater accessibility. Digital reenactments and multimedia adaptations of candidates' or officials' written materials make political information more engaging, understandable, and accessible, especially to voters with disabilities. Meanwhile, voters and political commentators also often use "audio or visual media that is digitally created or modified," § 20512(g), to express a range of messages—not only the messages just discussed, but also parody, satire, and the like.

B. Social media websites engage in content moderation—that is, determining whether and how to disseminate speech—including by regulating the dissemination of manipulated media.

That is not to say, however, that social media websites allow all content from all users. Rather, social media websites engage in extensive efforts to moderate the vast amounts of content that users create and post. Different websites approach moderation in varied ways. But all NetChoice member websites (and those regulated by AB2655) use content-moderation policies created by humans and designed to effectuate websites' bespoke editorial policies. Some websites rely more heavily on algorithms, which are programmed by humans, to identify, *e.g.*, misleading posts. Others prioritize human judgment for evaluating sensitive topics, while using algorithms

secondarily. Many websites also use features like user-generated flags, community notes, fact-check labels, and content warnings to further improve moderation accuracy. Although methods differ, these combined efforts effectively limit the dissemination of content those websites consider harmful or misleading—while protecting speech that complies with the websites' policies.

In particular, many social media websites' moderation policies specifically address misleading manipulations, deceptive content, or media that falsely portrays public figures.[3] At the same time, websites take care to avoid restricting political commentary, satire, and other beneficial, clearly labeled, or educational media. For example, videos and images that clarify candidate statements, facilitate accessibility, or aid voter understanding remain widely available. By balancing these competing considerations, social media websites protect users from what the websites regard as genuinely harmful misinformation while safeguarding the free exchange of valuable political content.

C. Content moderation can be difficult, and thus vague governmental demands to moderate content can chill the dissemination of speech.

Content moderation can be difficult, for multiple reasons. As an initial matter, content moderation requires websites to identify and manage content across diverse forms of media. Each medium presents its own challenge. For instance, as compared to text, it is especially time- and resource-intensive to moderate content containing images, audio, and video. These media types require advanced tools capable of analyzing visual and auditory elements. For example, automated systems must identify subtle manipulations, such as altered voices, synthesized imagery, or edited footage. Unlike text—which can be scanned for specific keywords or phrases—images and videos must be analyzed frame by frame, while audio requires transcription and contextual interpretation. The risk of error is higher because these systems must discern between legitimate alterations, such as filters or sound enhancements, and deceptive manipulations like deepfakes.

[3] *See, e.g.*, Meta Transparency Center, Misinformation, https://tinyurl.com/5fk3bv7v; Nextdoor, Nextdoor's misinformation policy, https://tinyurl.com/taxu7hhn; Pinterest, Community guidelines, https://perma.cc/J6JR-L6CY; X Help Center, Civic integrity policy, https://tinyurl.com/2fcunxve; Yahoo!, Yahoo Community Guidelines, https://perma.cc/YK6M-E9WB; YouTube, Elections misinformation policies, https://perma.cc/WFH5-STCV.

In addition to varying forms of media, content can be in different languages or reflect different cultures. So content that could appear to violate a website's content-moderation policies in one context may be innocuous in another. For example, many common hand gestures have both positive and negative connotations depending on cultural context. *See* David Anderson et al., *5 everyday hand gestures that can get you in serious trouble outside the US*, Business Insider (Jan. 5, 2019), https://perma.cc/8ZWT-JP5F. As a result, websites' content-moderation policies must account for regional variations and nuances.

The sheer scale of social media websites compounds these challenges. Many regulated websites disseminate billions of pieces of user-generated content daily. Automated tools are often essential for websites to ensure conformance with their policies. But over-reliance on automated tools can lead to both over-enforcement *and* under-enforcement if the tools lack the sophistication to interpret complex nuances. And those who wish to spread violative content constantly innovate their methods to evade detection. These tactics, such as minor alterations to deepfakes or using obscure languages and codes, force social media websites to continually adapt their moderation systems. This creates a dynamic and resource-intensive battle to define and enforce content-moderation standards.

D. AB2655 will make content moderation more difficult on regulated websites.

AB2655 intensifies content moderation's inherent difficulties in two key ways.

First, content moderation of political content can be especially difficult because it often involves sensitive judgments about context and intent. What constitutes deceptive or harmful material is frequently disputed, as political messaging can blur the lines between exaggeration and satire versus outright manipulation. Political content often circulates rapidly, amplifying its impact before websites can review it. Websites must also contend with the risk of perceived bias, as moderation decisions may appear to favor one political ideology or candidate over another, even when made in good faith. This perception can lead to public backlash and loss of user trust. The stakes are particularly high in election contexts, where errors in moderating political content can themselves undermine confidence in the democratic process or stoke division.

Second, the Act's requirements for how websites must implement their content moderation practices are vague. For example, the Act's overriding definition of "materially deceptive content" sets unclear boundaries between regulated and unregulated speech. § 20512(i)(1). The line between content that "would falsely appear to a reasonable person to be an authentic record of the content depicted in the media" versus content that "contains only minor modifications that do not significantly change the perceived contents or meaning of the content" is unclear. *Id.* The Act's exceptions for "satire" and "parody" only heighten this confusion. § 20519. What constitutes satire can often be in the eye of the beholder. Yet the Act would require covered websites to make those distinctions under the threat of guessing incorrectly what Defendants' after-the-fact judgments might be. In addition, it is not clear how social media websites can reliably identify each piece of content that is "reasonably likely" to "undermine confidence" in an election's outcome or "harm the reputation or electoral prospects of a candidate." § 20513(a)(2). Yet those terms are central to determining whether content must be removed or labeled. *See id.*; § 20514(a)(2)(A).

II. AB2655's requirements violate the First Amendment by forcing websites to moderate content according to California's preferences.

A. AB2655's requirements to remove certain political speech violate the First Amendment.

The Act's removal requirements violate core First Amendment protections for political speech and cannot satisfy any form of heightened scrutiny. *See* § 20513 (a "large online platform shall develop and implement procedures for the use of state-of-the-art techniques to identify and remove materially deceptive content," as defined by AB2655). This provision imposes content-based and viewpoint-based requirements for private websites to remove political speech from their services. Whatever interest the government may assert is already amply served by the opportunity for counter-speech and social media websites' existing content moderation. That is why this Court has preliminarily enjoined a materially similar law. *Kohls*, 2024 WL 4374134, at *6. And it is why other courts have enjoined other States' requirements to remove content-based categories of speech from their services. *E.g.*, *Comput. & Commc'ns Indus. Ass'n & NetChoice, LLC v. Paxton*, 747 F. Supp. 3d 1011, 1042-43 (W.D. Tex. 2024) ("*CCIA*"); *Fitch*, 2024 WL 3276409, at *17.

The First Amendment prohibits state laws that "grant the State virtually unfettered discretion to enforce as law a variety of subjective content policies." *Bonta II*, 2025 WL 807961, at *23.

1. The Act requires social media websites to remove content-based and viewpoint-based core political speech from their services, thus triggering strict scrutiny. States cannot regulate "the content choices the major platforms make . . . free of the First Amendment's restraints." *Moody*, 603 U.S. at 726-27. Nor may the government command private websites to remove speech that the government cannot regulate directly. *See Nat'l Rifle Ass'n of Am. v. Vullo*, 602 U.S. 175, 190 (2024). The Act's removal requirements trigger First Amendment scrutiny because they "deputize[] covered businesses into serving as censors for the State." *Bonta*, 113 F.4th at 1118.

The removal requirements are content-based because they target content about or relating to political "candidate[s]," "elections officials . . . in connection with the performance of their elections-related duties," and "elected official[s]" § 20513(a)(2). These regulated topics plainly "appl[y] to particular speech because of the topic discussed or the idea or message expressed." *City of Austin v. Reagan Nat'l Advert. of Austin, LLC*, 596 U.S. 61, 69 (2022) (citation omitted); *see Kohls*, 2024 WL 4374134, at *4 (Act "specifically targets speech within political or electoral content pertaining to candidates, electoral officials, and other election communication, making it a content-based regulation that seeks to limit public discourse"). Laws restricting access to or regulating speech based on content are subject to strict scrutiny. *Reed v. Town of Gilbert*, 576 U.S. 155, 163-64 (2015).

These removal requirements are also viewpoint-based because they only affect content that "is reasonably likely to *harm* the reputation or electoral prospects of a candidate" or "reasonably likely to falsely *undermine confidence* in the outcome of one or more election contests." § 20513(a)(2) (emphases added). False content that benefits a candidate's reputation or inspires confidence in elections is left unregulated. Likewise, the Act excludes "satire" and "parody." § 20519(c). "When the government targets not subject matter, but particular views taken by speakers on a subject, the violation of the First Amendment is all the more blatant." *Rosenberger v. Rector & Visitors of Univ. of Va.*, 515 U.S. 819, 829 (1995). "Viewpoint discrimination is thus an

egregious form of content discrimination." *Id.* And "[t]he government must abstain from regulating speech when the specific motivating ideology or the opinion or perspective of the speaker is the rationale for the restriction." *Id.*

2. Because these provisions target political speech and discriminate based on content, speaker, and viewpoint, they must satisfy strict scrutiny. *See Reed*, 576 U.S. at 163-64. They cannot. *See, e.g.*, *Kohls*, 2024 WL 4374134, at *4-5. The government has no legitimate, much less compelling, governmental interest in demanding that social media websites remove political speech from their services. *See, e.g.*, *NetChoice, LLC v. Reyes*, 748 F. Supp. 3d 1105, 1124-25 (D. Utah 2024). The State cannot "specifically identify" how the Act's scope responds to "an actual problem in need of solving" by *government* mandate. *Brown v. Ent. Merchs. Ass'n*, 564 U.S. 786, 799 (2011) (cleaned up). That is because the State lacks a legitimate interest in legislating truth or falsity, outside of narrow categories of unprotected speech that the Act does not regulate. *See United States v. Alvarez*, 567 U.S. 709, 721 (2012). "It is essential to a healthy democracy that 'debate on public issues [] be uninhibited, robust, and wide-open' which may create a necessary sacrifice that such dialogue 'include[s] vehement, caustic, and sometimes unpleasantly sharp attacks on government and public officials.'" *Kohls*, 2024 WL 4374134, at *4 (alterations in original) (quoting *N.Y. Times Co. v. Sullivan*, 376 U.S. 254, 270 (1964)).

Nor are § 20513's removal requirements "the least restrictive means" of furthering any sufficient governmental interest. *Ams. for Prosperity Found. v. Bonta*, 594 U.S. 595, 607 (2021); *see Kohls*, 2024 WL 4374134, at *1. Social media websites themselves already provide at least two less restrictive means. First, Defendant "cannot show[] why counterspeech would not suffice to achieve its interest." *Alvarez*, 567 U.S. at 726; *see Whitney v. California*, 274 U.S. 357, 377 (1927) ("If there be time to expose through discussion the falsehood and fallacies, to avert the evil by the processes of education, the remedy to be applied is more speech, not enforced silence."). Social media websites provide the means for the public to engage in a vast amount of counterspeech on their services. *See supra* pp.3-5. Promoting *more* speech on these services furthers First Amendment principles. Second, even if the government somehow could assert some legitimate interest in *less speech*, NetChoice members engage in content moderation addressing much of the

same content the State has targeted for regulation. The Supreme Court has held that "voluntary," industry-led self-regulatory efforts are less restrictive means than government intervention. *Brown*, 564 U.S. at 803. And even if the State deems those private efforts insufficient, "government does not have a compelling interest in each marginal percentage point by which its goals are advanced." *Id.* at 803 n.9.

In addition to these private alternatives, separate California law already addresses any *un-protected* speech that the State may legitimately regulate: "Other statutory causes of action such as privacy torts, copyright infringement, or defamation already provide recourse to public figures or private individuals whose reputations may be afflicted by artificially altered depictions[.]" *Kohls*, 2024 WL 4374134, at *5. "[P]rotected speech may [not] be banned as a means to ban un-protected speech." *Ashcroft v. Free Speech Coal.*, 535 U.S. 234, 255 (2002). Otherwise, "'[t]here are broad areas in which any attempt by the state to penalize purportedly false speech would present a grave and unacceptable danger of suppressing truthful speech.'" *Kohls*, 2024 WL 4374134, at *5 (alteration in original) (quoting *Alvarez*, 567 U.S. at 731 (Breyer, J., concurring in the judgment)).

Regardless, the Act's removal requirements are both "seriously underinclusive" and "seriously overinclusive." *Brown*, 564 U.S. at 805. The "overbreadth in achieving one goal is not cured by the underbreadth in achieving the other." *Id.* The Act is underinclusive because it regulates only disfavored websites, leaving unregulated myriad potential sources of the kinds of content that the State seeks to regulate. "Materially deceptive" political content (as defined by AB2655) can be found on countless websites that fall short of the Act's arbitrary 1-million-California-user threshold. § 20512(h). Yet the Act would leave those websites untouched. The Act also expressly allows "online newspaper[s]" and other Internet websites to disseminate the same speech that must be removed from social media websites, § 20519, many of which count such "online newspaper[s]" as registered users. Although these exempted sources must make disclosures about the authenticity of the content, they are not required to remove the content—as social media websites are. At the same time, the requirements for social media are vastly overinclusive. These websites already have their own content-moderation policies that address much of the content that the Act targets. *See*

supra p.5. These policies are an existing private alternative to the Act's overinclusive, one-size-fits-all governmental mandate. *See United States v. Playboy Ent. Grp.*, 529 U.S. 803, 826 (2000).

In addition to the Act's over- and underinclusive range of regulated media, the Act's definitions of prohibited speech are both overbroad and underinclusive. It is unclear why California requires removal of a subset of political speech, but not other forms of manipulated media.

B. The labeling requirements violate the First Amendment.

The Act's compelled-speech labeling requirements trigger First Amendment strict scrutiny as content-based and viewpoint-based regulations of speech. § 20514 (a "large online platform shall develop and implement procedures for the use of state-of-the-art techniques to identify materially deceptive content," as defined by AB2655). Compelled speech violates the First Amendment: The "freedom of speech includes both the right to speak freely and the right to refrain from speaking at all." *Janus v. Am. Fed'n of State, Cnty., & Mun. Emps., Council 31*, 585 U.S. 878, 892 (2018) (cleaned up). This freedom from laws compelling speech applies "not only to expressions of value, opinion, or endorsement, but equally to statements of fact." *Hurley v. Irish-Am. Gay, Lesbian & Bisexual Grp. of Bos.*, 515 U.S. 557, 573 (1995).

1. Section 20514 compels speech in two ways that require regulated services to "opin[e] on whether and how certain controversial categories of content should be moderated." *X Corp. v. Bonta*, 116 F.4th 888, 901 (9th Cir. 2024).

First, Section 20514 requires websites to label content by appending the content with the State's message: "This [video, audio, or image] has been manipulated and is not authentic." § 20514(c). In other words, the Act compels social media websites to speak about the "authentic[ity]" of content in words specifically identified by the State. *Id.*

Second, Section 20514 mandates that the "label . . . shall permit users to click or tap on it for additional explanation about the materially deceptive content in an easy-to-understand format." § 20514(d). Thus, websites ostensibly must separately explain *why* they concluded the content was "not authentic" or how the content "has been manipulated." § 20514(c). And critically, whether explicitly or implicitly, these requirements also compel websites to speak about whether content

is "satire" or "parody." § 20519. After all, if a social media website has labeled content pursuant to the Act, it has necessarily determined that the content is not exempted "parody" or "satire." *Id.*

2. These compelled-speech requirements trigger strict scrutiny for multiple reasons. Like Section 20513's removal requirements, Section 20514's labeling provisions apply to content about or relating to political "candidate[s]," "elections official[s] . . . in connection with the performance of their elections-related duties," and "elected official[s]" § 20514(a)(2) (incorporating § 20513(a)(2)). These compelled speech regulations are thus content-based. They are also viewpoint-based: They only apply to content "reasonably likely to *harm* the reputation or electoral prospects of a candidate" or "reasonably likely to falsely *undermine confidence* in the outcome of one or more election contests" and not "materially deceptive content" that may burnish reputations or foster confidence positive content. § 20513(a)(2) (emphases added). Even assuming those lines can be easily drawn at the government's behest (they cannot be), they are nevertheless viewpoint-based distinctions.

The labeling requirements apply to two more content-based categories of speech: (1) "advertisement[s]," which means content with "the purpose of supporting or opposing a candidate for elective office," § 20512(a); and (2) "election communication[s]," which include content that "concerns," *e.g.*, "candidate[s]," "[v]oting or refraining from voting," "canvass of the vote," "[v]oting machines, ballots, voting sites," and "[p]roceedings or processes of the electoral college," § 20512(e). Even if the labeling requirements did not regulate speech based on content and viewpoint, compelled-speech requirements always trigger strict scrutiny because they alter the content of speech. *See Riley v. Nat'l Fed'n of the Blind of N.C., Inc.*, 487 U.S. 781, 795 (1988); *see Kohls*, 2024 WL 4374134, at *5 (similar "disclosure requirement forces parodists and satirists to speak a particular message that they would not otherwise speak, which constitutes compelled speech that dilutes their message" (cleaned up)).

The commercial-speech doctrine, or some other lesser form of First Amendment scrutiny, does not apply here. Opining on the *authenticity* of political speech does not "propose a commercial transaction." *Bonta*, 113 F.4th at 1119 (quoting *Bolger v. Youngs Drug Prods. Corp.*, 463 U.S. 60, 66 (1983)); *see X Corp.*, 116 F.4th at 900 (similar). That is why this Court concluded that

similar labels were not commercial speech. *Kohls*, 2024 WL 4374134, at *5. This is true even though the Act requires labels for "advertisement[s]." § 20514(a)(2)(B). As defined by the Act, "advertisement[s]" are limited to *political speech* "*with the purpose of supporting or opposing a candidate for elective office*." § 20512(a) (emphasis added).

3. The labeling requirements fail strict scrutiny and any other form of First Amendment scrutiny. All the tailoring flaws that apply to Section 20513's removal requirements apply to Section 20514's labeling requirements. The Act's heavy-handed speech regulations are not properly tailored to achieve whatever sufficient interest the government might have. The government cannot demonstrate why the potential for counter-speech and websites' existing content moderation are insufficient to address the State's concerns. Nor is it clear why websites must make the State's prescriptive and burdensome disclosures, but a (1) "regularly published online newspaper, magazine, or other periodical" must only make "a clear disclosure"; and (2) "broadcasting station" must only "clearly acknowledge[] through content or a disclosure, in a manner that can be easily heard or read." § 20519(a), (b)(1).

Regardless, the labeling requirements are "unduly burdensome" because they have the potential to "drown[] out" the message in the user-generated content. *Nat'l Inst. of Fam. & Life Advocs. v. Becerra*, 585 U.S. 755, 778 (2018). For example, the Act creates an incentive for websites to avoid liability by labeling satire and parody as "materially deceptive." But that labeling robs those formats of their value. Satire often works precisely because viewers initially mistake it for genuine content. Its effectiveness comes from the moment audiences discover their error after reacting sincerely. Labeling content "materially deceptive" from the outset removes that moment of realization. It also shapes how they interpret the message, as they approach it already aware of its intended irony or humor. Such mandated labels thus undercut the purpose and value of satire itself. The labeling requirements are also unduly burdensome user-generated content that is in a non-English language. For content that meets that description, websites must include two labels: one in the content's original language, and one in English. § 20517.

In addition to the effect on the user-generated speech on the services, these requirements are burdensome to implement and manage. As the Plaintiffs have explained, the requirement to

label content within 72 hours of a report, § 20513(b), will be costly and difficult to implement. It does not matter that some social media websites already append their own labels to a similar range of content as regulated under the Act. Those services label content using their own words, according to their own editorial policies, *not* the State's preferred message. And they operate according to the companies' existing capacities and timelines—not according to the Act's quick turnaround times.

C. The reporting-and-response requirements violate the First Amendment.

The Act's requirements for websites to establish a reporting-and-response procedure for the prohibited categories of political speech—and respond to users within 36 hours—violate the First Amendment. § 20515(a). These requirements purport to direct websites' content moderation by effectuating the removal and labeling requirements. They also compel speech by requiring websites to "describe[e] any action taken or not taken . . . with respect to the content." § 20515(a). The 36 hours that the Act mandates for responses will burden the dissemination of speech and require snap decisions—providing incentives to be overinclusive in restricting speech.

1. This reporting-and-response requirement is a content-based and viewpoint-based regulation of speech for the same reasons as the removal and labeling requirements: It allows people to report content-based and viewpoint-based categories of speech they think "should be removed . . . or labeled." § 20515(a); *see supra* p.9. In turn, the Act intends for these reports to direct social media websites' content moderation, as only reported content must either be removed or labeled. §§ 20513(a)(1), 20514(a)(1). Thus, this provision "deputizes covered businesses into serving as censors for the State." *Bonta*, 113 F.4th at 1118.

This provision also compels speech—triggering strict scrutiny in yet another way. *See Riley*, 487 U.S. at 795. It requires websites to "describe[e] any action taken or not taken by the large online platform with respect to the content." § 20515(a). Whenever someone identifies content through a report, the website must explain why it will (or will not) take a particular content-moderation action. Accordingly, much like the labeling requirement above, this provision mandates that websites must "opin[e] on whether and how certain controversial categories of content should

be moderated." *X Corp.*, 116 F.4th at 901. And the "First Amendment's guarantee of freedom of speech" equally prohibits "compelled speech and compelled silence." *Id.* at 900 (cleaned up).

2. These provisions cannot satisfy any form of heightened scrutiny. The only governmental interest this provision serves is to effectuate the unconstitutional removal and labeling requirements. Because those are unlawful, so too is the reporting-and-response requirement. Moreover, the 36-hour period to respond to these requests will be quite burdensome, as Plaintiff has explained. They will require social media websites to develop mechanisms to field numerous reports about even more content and then come to a considered decision in less than two days.

III. AB2655's removal and labeling requirements are preempted by 47 U.S.C. § 230.

Independent of the First Amendment violations, 47 U.S.C. § 230 preempts the Act's removal and labeling requirements (§§ 20513-14). These requirements command social media websites to conform their private content moderation to the State's standards. As the Ninth Circuit has explained, "Section 230 prohibits holding companies responsible for moderating or failing to moderate content." *Est. of Bride v. Yolo Techs., Inc.*, 112 F.4th 1168, 1182 (9th Cir. 2024).

Section 230 protects a broad range of "publisher[s'] traditional editorial functions." *Barnes v. Yahoo!, Inc.*, 570 F.3d 1096, 1102 (9th Cir. 2009) (citation omitted). That includes "reviewing, editing, and deciding whether to publish or to withdraw from publication third-party content." *Id.*; *see* 47 U.S.C. § 230(f)(4) (protecting websites' right to "filter, screen, allow, [] disallow," "pick, choose," "display," "organize," and "reorganize . . . content"). Specifically, no "interactive computer service shall be treated as the publisher or speaker of any information provided by" someone else. 47 U.S.C. § 230(c)(1). And websites cannot be held liable on account of "any action voluntarily taken in good faith to restrict access to" speech they or their users "consider[] to be . . . objectionable." *Id.* § 230(c)(2)(A). Congress preempted "inconsistent" state laws, protecting websites from "cause[s] of action" and "liability." *Id.* § 230(e)(3).

AB2655, however, provides a "cause of action" and potential liability for alleged failures to moderate content as the State requires. Social media websites must "develop and implement procedures for the use of state-of-the-art techniques to *identify materially deceptive content* and" "remove" and "label[]" such content. §§ 20513-14. That will require these websites to monitor

content and either remove or label content meeting the Act's requirements. Section 230 preempts all of this, as binding Ninth Circuit precedent makes clear.

"[S]ection 230 protects from liability 'any activity that can be boiled down to deciding whether to exclude material that third parties seek to post online.'" *Barnes*, 570 at 1103 (quoting *Fair Hous. Council of San Fernando Valley v. Roommates.com, LLC*, 521 F.3d 1157, 1170-71 (9th Cir. 2008) (en banc)). And because of the requirement to "identify" content, the Act "necessarily require[s]" social media websites "to monitor third-party content" so that users are not exposed to prohibited content. *HomeAway.com, Inc. v. City of Santa Monica*, 918 F.3d 676, 682 (9th Cir. 2019). That is why the Western District of Texas recently enjoined a state requirement to monitor for, and remove, content-based categories of speech. *CCIA*, 747 F. Supp. 3d at 1042-43.

Section 230 likewise bars any requirement that websites alternatively apply a label warning the public of such prohibited content. *E.g.*, *Yolo*, 112 F.4th at 1180 ("Similarly, the failure to warn claim faults YOLO for not mitigating, in some way, the harmful effects of the harassing and bullying content. This is essentially faulting YOLO for not moderating content in some way, whether through deletion, change, or suppression."); *Doe v. Grindr Inc.*, 128 F.4th 1148, 1154 (9th Cir. 2025) ("Grindr's role as a publisher of third-party content does not give it a duty to warn users of a general possibility of harm resulting from the App." (cleaned up)); *Wozniak v. YouTube, LLC*, 100 Cal. App. 5th 893, 914-15 (2024) ("Plaintiffs' argument would allow essentially every state cause of action otherwise immunized by section 230 to be pleaded as a failure to warn of such information published by a defendant.").

AB2655 thus directly conflicts with Section 230, because it imposes liability on social media websites that fail to censor or label third-party content exactly as the State demands. By requiring social media websites to remove or label speech according to state-imposed standards, AB2655 interferes with those websites' protected editorial judgments. The whole point of Section 230 is to broadly protect websites from such state attempts to dictate content moderation policies or impose liability for third-party speech. Allowing California to enforce AB2655 would essentially nullify the protections Congress created in Section 230. Section 230 therefore independently preempts AB2655's removal and labeling mandates.

CONCLUSION

This Court should grant Plaintiffs' motion for summary judgment and permanently enjoin enforcement of AB2655 against all regulated websites.

DATED: March 27, 2025

Bradley A. Benbrook
Bradley A. Benbrook (SBN 177786)
Stephen M. Duvernay (SBN 250957)
BENBROOK LAW GROUP, PC
701 University Avenue, Suite 106
Sacramento, CA 95825
Telephone: (916) 447-4900
brad@benbrooklawgroup.com
steve@benbrooklawgroup.com

Steven P. Lehotsky*
Scott A. Keller*
Jeremy Evan Maltz*
LEHOTSKY KELLER COHN LLP
200 Massachusetts Avenue, NW, Suite 700
Washington, DC 20001
(512) 693-8350
steve@lkcfirm.com
scott@lkcfirm.com
jeremy@lkcfirm.com

Joshua P. Morrow*
LEHOTSKY KELLER COHN LLP
408 W. 11th Street, 5th Floor
Austin, TX 78701
(512) 693-8350
josh@lkcfirm.com

* *Pro hac vice* forthcoming.

Attorneys for Amicus Curiae NetChoice